

PENERAPAN *SEASONAL GENERALIZED SPACE TIME AUTOREGRESSIVE SEEMINGLY UNRELATED REGRESSION (SGSTAR SUR)* PADA PERAMALAN HASIL PRODUKSI PADI

Leni Pamularsih¹, Mustafid², Abdul Hoyyi³

^{1, 2, 3} Departemen Statistika Fakultas Sains dan Matematika Universitas Diponegoro

email: lenipamularsih@gmail.com

ABSTRACT

Ordinary Least Square (OLS) is general method to estimate Generalized Space Time Autoregressive (GSTAR) parameters. Parameter estimation by using OLS for GSTAR model with correlated residuals between equations will produce inefficient estimators. The method that appropriate to estimate the parameter model with correlated residuals between equations is Generalized Least Square (GLS), which is usually used in Seemingly Unrelated Regression (SUR). This research aims to build the seasonal GSTAR SUR model as model of rice yield forecasting in three locations by using the best weighting. Weights used are binary weights, inverse distance and normalization of cross correlation. Data which used in this research are the data of rice yield per quarter in three districts in Central Java, namely Banyumas, Cilacap and Kebumen. The data from the period of January 1981 to December 2014 as training data and the period of January 2015 to December 2018 as validation data. The resulting is a model that has a seasonal effect with the autoregressive order and the spasial order limited to 1 so the model formed is SGSTAR (4₁)-I(1)(1)³. The best model produced is the SGSTAR SUR (4₁)-I(1)(1)³ model with inverse distance weighting because it fulfills both assumptions, residuals white noise and residuals normally multivariate distribution. Additionally, it has the smallest MAPE value when compared the other weighting, that is 20%. This MAPE value indicates that the accuracy rate of forecast is accurate.

Keywords: Rice yield, Seasonal, GSTAR, SUR.

1. PENDAHULUAN

Indonesia adalah negara agraris memiliki potensi besar di bidang pertanian. Namun, sebagai negara agraris tidak menjadikan Indonesia mudah untuk memenuhi kebutuhan masyarakat Indonesia. Salah satunya adalah luas lahan yang tidak memadai. Menurut Kepala BPS RI Suharyanto dalam berita economy.okezone.com mengatakan bahwa terjadi penurunan luas lahan baku lahan pertanian di Indonesia menjadi 7,1 juta hektare pada tahun 2018 dibandingkan dengan data sensus 2013 luas bahan baku lahan pertanian mencapai 7,75 hektare. Dengan luas lahan tersebut diharapkan mampu menghasilkan komoditas pangan yaitu padi yang mampu memenuhi kebutuhan masyarakat Indonesia. Namun, pada kenyataannya produksi beras kerap kali tidak mencukupi kebutuhan dalam negeri, sehingga tidak jarang Indonesia melakukan impor beras. Berdasarkan data dari BPS, pada tahun 2018 Indonesia melakukan impor beras dari beberapa negara seperti Thailand, Vietnam, India dan lain-lain. Kondisi tersebut berbanding terbalik dengan kondisi produksi beras di Provinsi Jawa Tengah yang mengalami surplus. Pada tahun 2018 Provinsi Jawa Tengah mengalami surplus produksi padi mencapai 1,78 juta ton dan pada tahun 2019 mengalami surplus 3,7 juta ton.

Atas dasar uraian tersebut maka perlu dilakukan peramalan hasil produksi padi di beberapa kabupaten di Jawa Tengah yang mampu menghasilkan produksi padi dalam jumlah

yang banyak yaitu Kabupaten Banyumas, Cilacap dan Kebumen. Hasil prediksi dapat digunakan untuk mengukur kemampuan produksi padi di kabupaten tersebut dalam memenuhi kebutuhannya atau bahkan di tingkat provinsi. Namun diharapkan mampu menjadi swasembada pangan di tingkat nasional, sehingga ketergantungan terhadap impor beras ke Negara lain dapat berkurang.

Salah satu metode yang dapat digunakan untuk memodelkan dan meramalkan data yang tidak hanya memiliki keterkaitan waktu tetapi juga lokasi yaitu *Generalized Space Time Autoregressive* (GSTAR). Model GSTAR diperkenalkan oleh Borovkova *et al.* (2002). Nilai parameter dari model GSTAR diperbolehkan untuk lokasi yang heterogen (Borovkova, Lopuhaä dan Ruchjana, 2008). Untuk mengestimasi parameter dari model GSTAR dibutuhkan suatu metode. Metode yang biasa digunakan adalah *Ordinary Least Square* (OLS) (Borovkova, Lopuhaä dan Ruchjana, 2008). Metode OLS dalam model GSTAR menjadi tidak efisien jika residual saling berkorelasi antar persamaan maka untuk mengatasinya digunakan metode *Generalized Least Square* (GLS), yang biasa digunakan dalam *Seemingly Unrelated Regression* (SUR) (Zellner, 1962).

2. TINJAUAN PUSTAKA

2.1. Runtun Waktu Multivariat

Data runtun waktu yang memiliki variabel lebih dari satu dan pengamatan yang dilakukan secara serentak pada dua seri waktu atau lebih disebut sebagai runtun waktu multivariate. Jika suatu variabel terdiri dari N variabel $\{z_{1,t}\}, \dots, \{z_{N,t}\}$. Maka runtun waktu multivariat merupakan vektor yang berukuran $(N \times 1)$ dinotasikan dengan $\{Z_t\}$. Prosedur multivariate time series sama halnya seperti univariat time series yaitu dimulai dengan identifikasi model, estimasi parameter, cek diagnosa dan peramalan.

Analisis runtun waktu hal yang fundamental yaitu stasioneritas. Data dikatakan stasioner jika pola data berada pada kesetimbangan rata-rata (*mean*) dan varian yang konstan selama waktu tertentu. Dalam analisis univariat *time series*, deret waktu yang non-stasioner direduksi menjadi deret waktu yang stasioner dengan menggunakan transformasi dan *differencing*. Hal tersebut juga digunakan dalam multivariat *time series*, namun transformasi harus diterapkan ke series komponen secara individual karena tidak semua series komponen dapat direduksi menjadi stasioner dengan bentuk transformasi yang sama. Stasioneritas dalam varian dapat dilihat dari transformasi *box cox*. Jika nilai *rounded value* bernilai 1 maka data dikatakan stasioner dalam varian. Selain itu juga dapat dilihat dari nilai batas atas dan batas bawah λ dari suatu data time series mengandung nilai 1. Stasioneritas data pada multivariate *time series* dilihat dari MPCCF (*Matrix Partial Cross Correlation Function*), MCCF (*Matrix Cross Correlation Function*) dan *box-cox transformation* (Wei, 2006).

Matrix Cross Correlation Function (MCCF) digunakan untuk melihat stasioneritas data dalam mean. Jika vektor *time series* terdiri dari n observasi, yaitu Z_1, Z_2, \dots, Z_n maka didapatkan persamaan fungsi matriks korelasi untuk $\rho(k) = [\rho_{i,j}(k)]$ diestimasi oleh fungsi matriks korelasi sampelnya adalah sebagai berikut (Wei, 2006) :

$$\hat{\rho}(k) = [\hat{\rho}_{i,j}(k)] \quad (1)$$

$[\hat{\rho}_{i,j}(k)]$ merupakan korelasi silang sampel diantara komponen series ke- i dan ke- j di lag ke- k , yang dinyatakan dalam persamaan berikut :

$$\hat{\rho}_{i,j}(k) = \frac{\sum_{t=1}^{n-k} (Z_{i,t} - \bar{Z}_i)(Z_{j,t+k} - \bar{Z}_j)}{[\sum_{t=1}^n (Z_{i,t} - \bar{Z}_i)^2 \sum_{t=1}^n ((Z_{j,t} - \bar{Z}_j)^2)]^{1/2}} \quad (2)$$

Matriks korelasi silang sangat berguna untuk menentukan model *Moving Average* (MA). Tiao dan Box (1981) dalam Wei (2006), memperkenalkan metode yang mudah untuk merangkum korelasi sampel. Metode tersebut yaitu untuk mencatat simbol (+), (-) dan (.) dalam posisi (i, j) dari matriks korelasi sampel. Simbol (+) menunjukkan bahwa nilai $\hat{\rho}_{i,j}(k)$ lebih besar dari 2 kali *standar error* (korelasi positif). Simbol (-) menunjukkan nilai $\hat{\rho}_{i,j}(k)$ kurang dari -2 kali *standard error* (korelasi negatif). Sedangkan simbol (.) menunjukkan nilai $\hat{\rho}_{i,j}(k)$ berada diantara ± 2 kali *standard error* (tidak ada korelasi).

Matrix Partial Cross Correlation (MPCCF) digunakan untuk mengidentifikasi model AR(p). Selain itu, MPCCF digunakan untuk mengidentifikasi orde waktu dari GSTAR. $\mathcal{P}(k)$ sama dengan $\Phi_{k,k}$ dalam regresi linier multivariat.

Persamaan untuk fungsi matriks autokorelasi spasial didefinisikan sebagai berikut :

$$\mathcal{P}(k) = \begin{cases} \Gamma'(1)[\Gamma(0)]^{-1}, & k = 1 \\ \{\Gamma'(k) - \mathbf{c}'(k)[\mathbf{A}(k)]^{-1}\mathbf{b}(k)\}[\Gamma(0) - \mathbf{b}'(k)[\mathbf{A}(k)]^{-1}\mathbf{b}(k)]^{-1}, & k > 1 \end{cases} \quad (3)$$

Untuk $k \geq 2$, maka didapatkan :

$$\mathbf{A}(k) = \begin{pmatrix} \Gamma(0) & \Gamma(1) & \dots & \Gamma(k-2) \\ \Gamma(1) & \Gamma(0) & \dots & \Gamma(k-3) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(k-2) & \Gamma(k-3) & \dots & \Gamma(0) \end{pmatrix}, \mathbf{c}(s) = \begin{pmatrix} \Gamma(1) \\ \Gamma(2) \\ \vdots \\ \Gamma(k-1) \end{pmatrix}, \mathbf{b}(k) = \begin{pmatrix} \Gamma'(k-1) \\ \Gamma'(k-2) \\ \vdots \\ \Gamma'(1) \end{pmatrix}, \quad (4)$$

Jika $\mathcal{P}(k)$ mendefinisikan model dari suatu data merupakan vektor AR(p), maka

$$\mathcal{P}(k) = \begin{cases} \Phi_p, & k = p \\ \mathbf{0}, & k > p \end{cases} \quad (5)$$

2.2. Generalized Space Time Autoregressive (GSTAR)

Jika diberikan suatu observasi yang terdiri dari lokasi i dan waktu t atau dapat dinotasikan $Z_{i,t}$ merupakan suatu proses stokastik, $t=1,2,\dots,T$ dan $i=1,2,\dots,N$. Proses vektor dari model GSTAR $\mathbf{Z}_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{N,t}]'$, sehingga dimensi dari vektor tersebut yaitu $(N \times 1)$. Model GSTAR dengan orde autoregressive (p) dan orde spasial $\lambda_1, \dots, \lambda_p$ (GSTAR(p ; $\lambda_1, \dots, \lambda_p$)) di definisikan sebagai berikut (Borovkova, Lopuhaä dan Ruchjana, 2008):

$$\mathbf{Z}_t = \sum_{k=1}^p [\Phi_{k0} \mathbf{Z}_{t-k} + \sum_{\ell=1}^{\lambda_{p0}} \Phi_{k\ell} \mathbf{W}^{(\ell)} \mathbf{Z}_{t-k}] + \mathbf{e}_t \quad t=0, \pm 1, \pm 2, \dots \quad (6)$$

Dengan $\Phi_{k0} = \text{diag}(\phi_{10}^{(k)}, \dots, \phi_{N0}^{(k)})$ merupakan parameter autoregressive, $\Phi_{k\ell} = \text{diag}(\phi_{1\ell}^{(k)}, \dots, \phi_{N\ell}^{(k)})$ merupakan parameter spasial, \mathbf{e}_t merupakan vektor residual yang memenuhi asumsi identik, independen dan berdistribusi normal multivariat dengan *mean* nol dan $E(\mathbf{e}_i \mathbf{e}_j') = \sigma_{ij} \mathbf{I}_n$.

Pada penelitian ini digunakan bobot biner, invers jarak dan normalisasi korelasi silang.

2.3. *Seemingly Unrelated Regression (SUR)*

Menurut Olamide dan Adepoju (2013), model *Seemingly Unrelated Regression (SUR)* merupakan generalisasi dari model regresi linier yang terdiri dari beberapa persamaan regresi dengan *residual* antar pengamatan dalam satu persamaan tidak berkorelasi, tetapi *residual* antara persamaan yang satu dengan yang lain berkorelasi. Menurut Greene (2002), model SUR dapat diestimasi dengan menggunakan metode *Generalized Least Square (GLS)*. Estimasi parameter model SUR dengan metode GLS memerlukan invers dari matriks varian kovarian *residual* (Ω^{-1}). Penaksir koefisien tak bias β dengan metode GLS yaitu Olamide dan Adepoju (2013):

$$\hat{\beta} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y \quad (7)$$

Asumsi yang harus dipenuhi dalam model SUR yaitu $E(e) = 0$ dan $E(e'e) = \sigma_{ij} I_T$, untuk $i, j=1, 2, \dots, N$ (Dwivedi dan Srivastava, 1978).

2.4. *Pengujian Asumsi Residual*

Asumsi yang harus dipenuhi yaitu residual harus bersifat *white noise* dan terdistribusi normal (Pfeifer dan Deutsch, 1980). Uji *white noise* dilakukan dengan memodelkan ulang residual yang didapatkan dari pemodelan. Pada penelitian ini untuk menguji *white noise residual* dilakukan dengan melihat plot MCCF (*Matrix Cross Correlation Function*) atau menggunakan kriteria minimum AIC (Wei, 2006). Jika nilai AIC terkecil terletak pada AR(0) dan MA(0) maka *residual* bersifat *white noise*. Sedangkan pengujian asumsi multivariat normal residual secara formal dengan menggunakan Kolmogorov Smirnov Test yaitu dengan H_0 menyatakan residual berdistribusi teoritis dan hipotesis alternatif atau H_1 menyatakan residual tidak berdistribusi teoritis. Dengan statistik uji :

$$D = \text{Sup} |S(d_j^2) - F_0(d_j^2)| \quad (8)$$

dengan $S(d_j^2)$ merupakan proporsi jarak mahalnobis yang $\leq d_j^2$ dan $F_0(d_j^2)$ adalah fungsi peluang kumulatif dari distribusi hipotesis (*chi-square*). H_0 ditolak jika $D > W_{(1-\alpha)}$ atau nilai $p.value < \alpha$ dengan uji dua sisi, dengan $W_{(1-\alpha)}$ adalah nilai dari tabel Kolmogorov-Smirnov dengan kuantil $1-\alpha$.

2.5. *Kriteria Pemilihan Model Terbaik*

Kriteria pemilihan model peramalan terbaik digunakan kriteria MAPE (*Mean Absolute Percentage Error*). Misalkan A_t dan F_t menunjukkan nilai aktual dan nilai perkiraan pada titik data t . MAPE dapat didefinisikan sebagai berikut (Heeyoung Kim dan Sungil Kim, 2016):

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{A_t - F_t}{A_t} \right| \times 100\% \quad (9)$$

dengan N adalah banyaknya ramalan yang dilakukan. MAPE digunakan untuk mengetahui akurasi prediksi dari suatu model. Kriteria akurasi model berdasarkan nilai MAPE yaitu $MAPE < 10\%$ menunjukkan tingkat akurasi peramalan sangat akurat, $11\% < MAPE < 20\%$ memiliki tingkat akurasi peramalan akurat, $21\% < MAPE < 50\%$ memiliki tingkat akurasi cukup akurat dan $MAPE > 50\%$ menunjukkan tingkat akurasi peramalan tidak akurat (Abidin dan Jaffar, 2014).

3. METODOLOGI PENELITIAN

3.1. Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini yaitu data sekunder yang didapatkan dari Dinas Pertanian dan Perkebunan dan BPS Provinsi Jawa Tengah. Penelitian ini menggunakan data hasil produksi padi di tiga Kabupaten di Jawa Tengah yang meliputi Kabupaten Banyumas ($Z_{1,t}$), Kabupaten Cilacap ($Z_{2,t}$) dan Kabupaten Kebumen ($Z_{3,t}$) yang diambil per caturwulanan dari bulan Januari 1981 sampai dengan Desember 2018. Data pada periode Januari 1981-Desember 2014 digunakan sebagai data training (data *in-sample*) sedangkan data pada periode Januari 2015-Desember 2018 digunakan sebagai data validasi (data *out-sample*).

3.2. Langkah-langkah Analisis

Adapun tahapan analisis dalam penelitian ini yaitu sebagai berikut :

1. Melakukan identifikasi stasioner mean dan varian.
2. Melakukan identifikasi orde musiman dan non musiman serta orde model dugaan sementara dengan melihat MCCF, MPCCF serta nilai AIC minimum.
3. Membentuk model GSTAR OLS dengan langkah-langkah sebagai berikut :
 - a. Menetapkan nilai bobot lokasi (bobot biner, invers jarak dan normalisasi korelasi silang).
 - b. Melakukan penaksiran estimasi parameter dengan metode OLS per bobot lokasi.
 - c. Menguji signifikansi parameter model GSTAR OLS per bobot lokasi.
4. Menguji adanya korelasi residual antar lokasi atau tidak, apabila terdapat korelasi residual antar lokasi maka dibentuk model GSTAR SUR.
5. Membentuk model GSTAR SUR dengan langkah-langkah sebagai berikut:
 - a. Melakukan penaksiran parameter dari model GSTAR dengan model *Seemingly Unrelated Regression* (SUR) dengan metode *Generalized Least Square* (GLS).
 - b. Menguji signifikansi parameter model GSTAR SUR per bobot lokasi.
 - c. Menguji asumsi residual model GSTAR SUR untuk masing-masing bobot lokasi yaitu uji *white noise* dan normal multivariate.
7. Menentukan model GSTAR SUR yang terbaik diantara tiga jenis bobot yang digunakan berdasarkan nilai MAPE.
8. Melakukan peramalan dengan menggunakan model GSTAR SUR terbaik.

4. HASIL DAN PEMBAHASAN

4.1. Statistik Deskriptif

Data hasil produksi padi di tiga lokasi memiliki keterkaitan lokasi yang kuat karena memiliki nilai koefisien korelasi $\geq 0,9$ dan nilai *p.value* lebih kecil dari α ($\alpha = 5\%$) yang mengindikasikan ada korelasi antar lokasi. Korelasi yang sangat tinggi terjadi antara Kabupaten Banyumas dan Cilacap yaitu sebesar 0,93397.

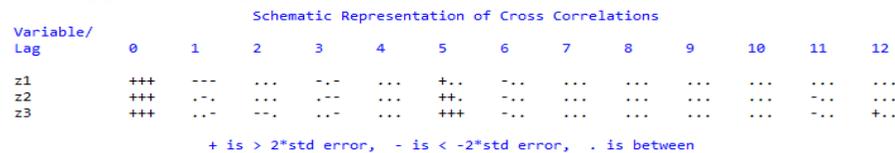
Rata-rata tertinggi yaitu di Kabupaten Cilacap dengan rata-rata hasil produksi padi 205.364 ton sedangkan rata-rata terendah di Kabupaten Banyumas dengan rata-rata 115.525 ton. Nilai standar deviasi menunjukkan heterogenan hasil produksi padi. Semakin besar nilai standar deviasi maka semakin besar sebarannya artinya data semakin bervariasi. Kabupaten yang memiliki heterogenan tertinggi yaitu Kabupaten Cilacap karena memiliki standar deviasi tertinggi yaitu 96.787 ton. Tiga lokasi memiliki pola yang sama yaitu rata-rata memiliki hasil produksi tertinggi pada periode 1 (bulan Januari- April) dan mulai turun pada periode 2

(bulan Mei-Agustus) dan pada periode 3 (bulan September-Desember) merupakan periode dengan rata-rata hasil produksi terendah untuk setiap kabupaten.

4.2. Stasioneritas

Stasioner dalam varian dilihat dari *box cox transformation* didapatkan nilai *rounded value* untuk data hasil produksi padi di kabupaten Banyumas, Cilacap dan Kebumen yaitu 0,50, 0,50 dan 0,50. Hal tersebut menunjukkan bahwa nilai *rounded value* $\neq 1$ yang mengindikasikan data belum stasioner dalam varian. Oleh karena itu, dilakukan transformasi dengan menggunakan bentuk transformasi $\sqrt{Z_t}$. Setelah ditransformasi didapatkan nilai *rounded value* = 1 untuk semua lokasi yang menunjukkan bahwa data sudah stasioner dalam varian. Untuk melihat stasioneritas dalam mean secara multivariate dapat dilihat dari skema plot MCCF. Dari skema plot MCCF data belum stasioner dalam *mean* pada tingkat level maupun *differencing* 1 terlihat bahwa pola yang terbentuk sama di antara keduanya dengan tanda (+) dan (-) masih banyak ditemukan dalam plot MCCF artinya ada korelasi positif ataupun negatif.

Dari plot MCCF baik tingkat level maupun *differencing* 1, perubahan tanda (+) dan (-) membentuk pola dengan tanda (+) akan muncul setiap kelipatan 3, hal tersebut mengindikasikan adanya pola musiman dengan periode 3. Oleh karena itu dilakukan *differencing* dengan *differencing* 1 dan 3, dengan *differencing* 1 merupakan *differencing* untuk non musiman dan *differencing* 3 untuk *differencing* musimannya. Plot MCCF hasil *differencing* dengan menggunakan *differencing* 1 dan 3 dapat dilihat pada Gambar 1. Dari Gambar 1 terlihat bahwa tanda (.) lebih banyak jika dibandingkan tanda (+) dan tanda (-). Hal tersebut mengindikasikan bahwa tidak adanya korelasi yang menunjukkan data sudah stasioner dalam *mean* dengan menggunakan *differencing* 1 dan 3.



Gambar 1. Skema Plot MCCF *differencing* 1 dan 3

Untuk memastikan pengujian stasioneran dalam *mean* yang maka dilakukan pengujian secara formal dengan menggunakan *Augmented Dickey Fuller Test* (ADF test). Dari pengujian *Augmented Dickey Fuller* (ADF) menunjukkan bahwa nilai *p.value* dari masing-masing lokasi yaitu 0,0000. Nilai *p.value* tersebut lebih kecil jika dibandingkan dengan $\alpha = 5\%$, yang artinya H_0 ditolak sehingga dapat disimpulkan bahwa data *differencing* 1 dan 3 telah stasioner dalam *mean*. Oleh karena itu, untuk analisis selanjutnya menggunakan data hasil transformasi dan *differencing* 1 dan 3.

4.3. Identifikasi Model

Identifikasi model GSTAR dapat dilihat dari skema plot MPCCF dari data yang sudah stasioner serta nilai AIC terkecil. Berdasarkan Tabel 1 hasil orde AR yang akan digunakan yaitu orde AR 4 MA 0, hal tersebut dikarenakan memiliki nilai AIC terkecil jika dibandingkan dengan nilai AIC pada lag-lag yang lain.

Tabel 1. Nilai Minimum *Akaike'a Information Criterion* (AIC)

Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
-----	------	------	------	------	------	------

AR 0	23.836142	23.352233	23.406028	23.328941	23.299526	23.331965
AR 1	23.035283	23.019852	22.989333	22.960486	22.980179	23.076943
AR 2	22.952540	23.019523	23.054461	22.654353	22.732383	22.874894
AR 3	22.632754	22.795312	22.890023	22.666606	22.825257	23.027706
AR 4	22.394523	22.648032	22.729694	22.63126	22.926829	22.987154
AR 5	22.537335	22.666372	22.723518	22.654401	22.91019	23.008315

Model GSTAR yang terbentuk yaitu orde autoregressive 4 dan untuk orde spasial dibatasi 1 serta memiliki efek *seasonal* sehingga dapat dituliskan model untuk pemodelan hasil produksi padi yaitu model *Seasonal* GSTAR $(4_1)-I(1)(1)^3$ atau SGSTAR $(4_1)-I(1)(1)^3$.

4.4. Estimasi Parameter dengan SGSTAR OLS

Untuk mengestimasi parameter model SGSTAR $(4_1)-I(1)(1)^3$ dapat digunakan metode *Ordinary Least Square* (OLS). Pada penelitian ini tidak dilakukan pemodelan ulang, jika terdapat estimasi parameter yang tidak signifikan. Hal tersebut didasarkan pada Kostenko dan Hyndman (2008) mengatakan bahwa parameter yang tidak signifikan dapat digunakan untuk peramalan, sehingga dalam penelitian ini parameter yang tidak signifikan tetap digunakan dalam peramalan. Berikut model SGSTAR OLS $(4_1)-I(1)(1)^3$ per masing-masing pembobot.

a. Bobot Biner

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -1,23727 & 0,12963 & 0,12963 \\ -0,00872 & -0,55684 & -0,00872 \\ 0,05090 & 0,05090 & -0,73373 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,74236 & 0,02090 & 0,02090 \\ -0,03669 & -0,45861 & -0,0367 \\ 0,071211 & 0,07121 & -0,59774 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,57690 & -0,01386 & -0,01386 \\ -0,10898 & -0,42721 & -0,10898 \\ 0,04905 & 0,04905 & -0,71462 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,22912 & -0,03593 & -0,03593 \\ -0,07241 & -0,25875 & -0,07241 \\ 0,05252 & 0,05252 & -0,44022 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

b. Bobot Invers Jarak

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -1,25238 & 0,167031 & 0,09294 \\ -0,10917 & -0,49700 & -0,04390 \\ 0,044999 & 0,03252 & -0,71586 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,77079 & 0,03892 & 0,02166 \\ -0,08246 & -0,43697 & -0,03316 \\ 0,08595 & 0,06211 & -0,59627 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,60167 & -0,00864 & -0,0048 \\ -0,11158 & -0,46211 & -0,04487 \\ 0,06876 & 0,049691 & -0,72464 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,26800 & -0,03130 & -0,01741 \\ -0,03566 & -0,31358 & -0,01434 \\ 0,07204 & 0,05206 & -0,44676 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

c. Bobot Normalisasi Korelasi Silang

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -1,23591 & 0,12776 & 0,13094 \\ 0,01273 & -0,58664 & 0,02503 \\ 0,04101 & 0,08701 & -0,76019 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,74102 & 0,02019 & 0,02070 \\ -0,01749 & -0,47077 & -0,03439 \\ 0,04023 & 0,08537 & -0,59677 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,57588 & -0,01401 & -0,01436 \\ -0,07189 & -0,42077 & -0,14133 \\ 0,02137 & 0,04534 & -0,69710 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,22749 & -0,03601 & -0,0369 \\ -0,05462 & -0,23937 & -0,10738 \\ 0,02321 & 0,04924 & -0,42679 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

Sebelum memodelkan GSTAR dengan menggunakan *Seemingly Unrelated Regression* (SUR) dilakukan pengujian korelasi residual terlebih dahulu. Jika residual antar persamaan saling berkorelasi maka dilanjutkan memodelkan dengan menggunakan GSTAR SUR. Jika nilai AIC minimum berada di AR(0) dan MA(0) maka tidak terdapat korelasi residual. Berdasarkan nilai AIC minimum untuk bobot biner, invers jarak dan normalisasi korelasi silang nilai minimumnya terdapat di AR(0) dan MA(3) yang artinya terdapat korelasi residual.

4.5. Estimasi Parameter dengan SGSTAR SUR

Sama halnya dengan SGSTAR OLS dengan SGSTAR SUR pun tidak dilakukan pemodelan ulang jika terdapat parameter yang tidak signifikan. Berikut model SGSTAR SUR (4_1) -I(1)(1)³ per masing-masing pembobot.

a. Bobot Biner

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -0,99544 & 0,07499 & 0,07499 \\ 0,13712 & -0,73341 & 0,13712 \\ 0,02630 & 0,02630 & -0,68734 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,68331 & 0,00819 & 0,00819 \\ 0,07824 & -0,58846 & 0,07824 \\ 0,02527 & 0,02527 & -0,49503 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,73035 & 0,01761 & 0,01761 \\ -0,04336 & -0,50713 & -0,04336 \\ -0,02002 & -0,02002 & -0,60055 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,36597 & -0,0092 & -0,0092 \\ 0,01912 & -0,37311 & 0,01912 \\ -0,04466 & -0,04466 & -0,27875 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

b. Bobot Invers Jarak

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -1,00051 & 0,09422 & 0,05243 \\ 0,15536 & -0,68535 & 0,06247 \\ 0,04937 & 0,03568 & -0,71901 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,69504 & 0,01592 & 0,00886 \\ 0,10712 & -0,57121 & 0,04307 \\ 0,04323 & 0,03124 & -0,5061 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,75455 & 0,03109 & 0,01730 \\ -0,06168 & -0,49614 & -0,0248 \\ -0,01612 & -0,01165 & -0,60524 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,38859 & -0,00091 & -0,00051 \\ 0,05611 & -0,38239 & 0,02256 \\ -0,04805 & -0,03473 & -0,27877 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

c. Bobot Normalisasi Korelasi Silang

$$\begin{bmatrix} \nabla Z_{1,t} \\ \nabla Z_{2,t} \\ \nabla Z_{3,t} \end{bmatrix} = \begin{bmatrix} -1,00690 & 0,07850 & 0,08045 \\ 0,08644 & -0,72971 & 0,16993 \\ 0,01879 & 0,03986 & -0,68817 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-1} \\ \nabla Z_{2,t-1} \\ \nabla Z_{3,t-1} \end{bmatrix} + \begin{bmatrix} -0,69072 & 0,00978 & 0,01003 \\ 0,04816 & -0,58799 & 0,09467 \\ 0,01217 & 0,02581 & -0,49031 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-2} \\ \nabla Z_{2,t-2} \\ \nabla Z_{3,t-2} \end{bmatrix} + \begin{bmatrix} -0,72407 & 0,01557 & 0,01595 \\ -0,02595 & -0,51307 & -0,05100 \\ -0,0126 & -0,02674 & -0,59607 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-3} \\ \nabla Z_{2,t-3} \\ \nabla Z_{3,t-3} \end{bmatrix} + \begin{bmatrix} -0,36303 & -0,0102 & -0,01046 \\ 0,00925 & -0,36904 & 0,01817 \\ -0,02551 & -0,05412 & -0,28147 \end{bmatrix} \begin{bmatrix} \nabla Z_{1,t-4} \\ \nabla Z_{2,t-4} \\ \nabla Z_{3,t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \end{bmatrix}$$

4.6. Pengujian Asumsi *White Noise Residual* dan Distribusi Normal Multivariat pada Model SGSTAR SUR (4_1) -I(1)(1)³

Pada pengujian *residual white noise* pada digunakan nilai AIC minimum. Dari ketiga pembobot yang digunakan yaitu bobot biner, invers jarak dan normalisasi korelasi silang nilai AIC minimum terdapat di AR(0) dan MA(0) sehingga dapat disimpulkan bahwa asumsi *residual white noise* terpenuhi. Sedangkan pengujian residual berdistribusi normal multivariate menggunakan Kolmogorov smirnov. Dari masing-masing pembobot memiliki nilai *p.value* > $\alpha = 5\%$ yang artinya H_0 diterima sehingga dapat disimpulkan bahwa data *residual* berdistribusi normal multivariat.

4.7. Pemilihan Model Terbaik

Dalam penelitian ini untuk memilih model terbaik digunakan nilai MAPE. Berikut Tabel 2 merupakan nilai MAPE per masing-masing bobot.

Tabel 2. Nilai MAPE pada Tiga Pembobot per Masing-masing Lokasi

Pembobot	Lokasi			Rata-rata MAPE
	Banyumas	Cilacap	Kebumen	
Biner	28,44531%	23,33561%	10,13870%	20,64247%
Invers Jarak	27,72966%	22,28300%	10,18533%	20,06600%
Normalisasi Korelasi Silang	29,03138%	23,54996%	10,54134%	21,04089%

Berdasarkan Tabel 2 menunjukkan bahwa nilai MAPE terkecil yaitu terdapat pada pembobot invers jarak dengan rata-rata MAPE yaitu $20,06600\% \cong 20\%$ yang menunjukkan bahwa tingkat akurasi peramalannya akurat. Model bobot invers jarak digunakan untuk memprediksi hasil produksi padi di tiga lokasi. Sebelum model tersebut digunakan untuk melakukan prediksi, model tersebut harus dikembalikan ke data asli dari *differencing* 1 dan 3 serta transformasi dengan transformasi yang digunakan yaitu $\sqrt{Z_t}$. Hasil prediksi produksi padi di tiga lokasi dari tahun 2019-2022 dapat dilihat pada Tabel 3.

Tabel 3. Hasil Peramalan Produksi Padi (Ton)

Tahun	Periode	Hasil Produksi Padi (Ton)		
		Banyumas	Cilacap	Kebumen
2019	Januari-April	107406	356346	311498
	Mei-Agustus	60132	248266	200564
	September-Desember	40212	193119	27837
2020	Januari-April	100083	365582	321490
	Mei-Agustus	54310	255750	209331
	September-Desember	35569	198252	30672
2021	Januari-April	92508	371177	329371
	Mei-Agustus	48890	261442	216005
	September-Desember	31225	204071	33670
2022	Januari-April	85425	379415	339440
	Mei-Agustus	43774	267656	223626
	September-Desember	27103	209075	36496

5. KESIMPULAN

Berdasarkan hasil dan pembahasan maka dapat disimpulkan bahwa model GSTAR yang terpilih didasarkan pada nilai AIC minimum yaitu terdapat pada orde waktu 4 dan orde spasial dibatasi satu dan karena memiliki efek *seasonal* sehingga dapat dituliskan model untuk produksi padi di tiga kabupaten Provinsi Jawa Tengah adalah model SGSTAR $(4)_1-I(1)(1)^3$ dengan penerapan tiga pembobot lokasi yaitu bobot biner, invers jarak dan normalisasi korelasi silang. Karena terdapat korelasi residual antar persamaan maka untuk mengatasinya digunakan

metode *Seemingly Unrelated Regression* (SUR). Model SGSTAR SUR terbaik yaitu model SGSTAR SUR $(4_1)-I(1)(1)^3$ dengan menggunakan bobot invers jarak. Hal tersebut dikarenakan pembobot invers jarak memenuhi asumsi *residual white noise* dan berdistribusi normal multivariate serta memiliki nilai MAPE terkecil diantara pembobot yang digunakan dengan rata-rata MAPE 20% yang menunjukkan bahwa model tersebut memiliki tingkat akurasi peramalan yang bagus atau akurat.

6. DAFTAR PUSTAKA

- Abidin, S.N.Z dan Jaffar, M.M. 2014. *Forecasting Share Prices of Small Size Companies in Bursa Malaysia Using Geometric Brownian Motion*. Applied Mathematics & Information Sciences An International Journal Vol.8, No. 1: Hal: 107-112.
- Borovkova, S. Lopuhaä, H.P., dan Ruchjana, B.N. 2008. *Consistency and Asymptotic Normality of Least Square Estimators in Generalized STAR Models*. Statistica Neerlandica Vol. 62, No. 4: Hal. 482-508.
- [BPS] Badan Pusat Statistik. 2020. Impor Beras Menurut Negara Asal Utama, 2000-2018. <https://www.bps.go.id/statictable/2014/09/08/1043/impor-beras-menurut-negara-asal-utama-2000-2018.html>. (28 Januari 2020)
- Dwivedi, T.D. and Srivastava, V.K. 1978. *Optimality of Least Squares in The Seemingly Unrelated Regression Equation Model*. Journal of Econometrics. Hal. 391-395.
- Greene, W.H. 2002. *Econometric Analysis*. Fifth Edition. New York University: Prentice-Hall.
- Kim, S dan Kim, H. 2016. *A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts*. International Journal of Forecasting. Hal. 669-679.
- Kostenko, A.V., dan Hyndman, R.J. *Forecasting Without Significance Test*. 2008
- Olamide, E.I. dan Adepoju, A.A. 2013. *Estimating Seemingly Unrelated Regression with First Order Autoregressive Disturbances*. Studies in Mathematical Science Vol. 6, No. 2, pp. 40-57.
- [Pemprov Jateng] Pemerintah Provinsi Jawa Tengah. 2020. Jelang Nataru Pangan Jateng Surplus, Siap Suplai Daerah Lain. <https://jatengprov.go.id/publik/jelang-nataru-pangan-jateng-surplus-siap-suplai-daerah-lain/>.(28 Januari 2020).
- Pfeifer, P.E. dan Deutsch, S.J. 1980. *A Three-Stage Iterative Procedure for Space Time Modelling*. Technometrics Vol. 22, No.1:Hal. 35-47.
- Wei, W.W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition. USA:Pearson Education, Inc.
- Zellner, A. 1962. *An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Agreggation Bias*. Journal of the American Statistical Association Vol. 57, No. 298, pp. 348-368